7 Lecture 7

7.1 Gauss elimination continued

Definition 7.1. A **system of linear equations** is of the form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \iff x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

In matrix notation, the equation above is written as $Ax = b$ where $A \in M_{m\times n}(F)$, $x \in F^n$ and $b \in F^m$. If $b = 0$, then the system is called **homogeneous**. If $b \neq 0$, the system is called **inhomogeneous**. We identify the system with the matrix A so we will refer to some columns as variables and variables as columns.

Definition 7.2. A matrix is in **row reduced echelon form (RREF)** if

(i) All zero rows, if any, are at the bottom.

(ii) In each nonzero row, the first nonzero entry (called a **pivot**) appears to the right of the pivot in the row above.

(iii) Each pivot equals 1.

(iv) Each pivot is the only nonzero entry in its column.

The non-pivot columns are called **free variables**.

Theorem 7.3. *For every matrix $A \in M_{m\times n}(F)$, one can apply elementary row operations to reduce A to a matrix R in RREF.*

(i) The number of pivot columns in R is $\operatorname{rank} A$. The columns of the original matrix A corresponding to the pivot columns form a basis for $\operatorname{im} A$.

(ii) The number of free variables is $n - \operatorname{rank} A = \operatorname{null} A$. To find a basis for $\ker A$, assign 1 to one free variable and 0 to the others, then solve the resulting system for the pivot variables, repeat this for each free variable. The resulting vectors form a basis for $\ker A$.

(iii) The nonzero rows of R form a basis for $\operatorname{im} A^t$.

(iv) The RREF of a matrix A is unique.

Remark 7.4. The pivot column must be a basis of $\operatorname{im} A$ but a basis of $\operatorname{im} A$ does not have to be the pivot columns.

Definition 7.5. The **augmented matrix** $[A \mid b]$ is the matrix obtained by appending the column vector $b \in F^m$ on the right to a matrix $A \in M_{m\times n}(F)$.

Theorem 7.6. *The system $Ax = b$ has a solution if and only if the last column in the RREF of $[A \mid b]$ is not a pivot column.*

If $Ax = b$ has a solution, then solutions of $Ax = b$ is given by solving the pivot variables in terms of the free variables.

We recall that Theorem 6.15 (ii) says $\operatorname{im} A = (\ker A^t)^{\perp}$. This is a dual space criterion which we can test solvability of $Ax = b$.

Theorem 7.7. *The equation $Ax = b$ has a solution if and only if for any $y \in F^m$ such that $A^t y = 0$ we have $y^t b = 0$.*

Proof. This is the content of Theorem 6.15 (ii). $\square$

Example 7.8. Let

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

We will compute:
(i) The kernel $\ker(A)$.
(ii) The image $\operatorname{im}(A^t) = \operatorname{im}(A)$.
(iii) The solution to the inhomogeneous system $Ax = b$.
(iv) Check $b \in \operatorname{im} A$ using the dual space criterion.
(v) A basis for $\operatorname{span}\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \bigcap \operatorname{span}\left\{\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}\right\}$.

We start with the augmented matrix and perform Gauss Elimination:

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 1 & 2 & 1 & 3 \\ -1 & 1 & 2 & 1 \end{array}\right)$$

Step 1: Swap R_1 and R_2 to get a leading 1 in the top-left corner.

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 2 & 1 & -1 & 2 \\ -1 & 1 & 2 & 1 \end{array}\right)$$

Step 2: Eliminate the entries below the pivot in column 1:

$$R_2 \leftarrow R_2 - 2R_1, \quad R_3 \leftarrow R_3 + R_1$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -3 & -3 & -4 \\ 0 & 3 & 3 & 4 \end{array}\right)$$

Step 3: Make the pivot in row 2 a 1:

$$R_2 \leftarrow -\frac{1}{3} R_2$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & \frac{4}{3} \\ 0 & 3 & 3 & 4 \end{array}\right)$$

Step 4: Eliminate above and below the pivot in column 2:

$$R_1 \leftarrow R_1 - 2R_2, \quad R_3 \leftarrow R_3 - 3R_2$$

$$\left(\begin{array}{ccc|c} 1 & 0 & -1 & \frac{1}{3} \\ 0 & 1 & 1 & \frac{4}{3} \\ 0 & 0 & 0 & 0 \end{array}\right)$$

(i) $Ax = 0$ is equivalent to

$$x_1 - x_3 = 0, \quad x_2 + x_3 = 0$$

So

$$\ker(A) = \left\{ \begin{pmatrix} x_3 \\ -x_3 \\ x_3 \end{pmatrix} : x_3 \in F \right\} = \operatorname{span} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

(ii) The first two columns of A are pivot columns, so

$$\operatorname{im} A = \operatorname{span} \left\{ \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}.$$

Since A is symmetric, $\operatorname{im} A = \operatorname{im} A^t$. So we also know that

$$\operatorname{im} A = \operatorname{im} A^t = \operatorname{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

The two bases are equivalent.

(iii) $Ax = b$ is equivalent to

$$x_1 - x_3 = \frac{1}{3}, \quad x_2 + x_3 = \frac{4}{3}$$

Thus the set of all solutions to $Ax = b$ are

$$\left\{ \begin{pmatrix} \frac{1}{3} + x_3 \\ \frac{4}{3} - x_3 \\ x_3 \end{pmatrix} : x_3 \in \mathbb{R} \right\} = \begin{pmatrix} \frac{1}{3} \\ \frac{4}{3} \\ 0 \end{pmatrix} + \operatorname{span} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

(iv) We test whether $Ax = b$ is consistent using the criterion:

$$Ax = b \iff y^t b = 0 \text{ for any } A^t y = 0$$

Since $A = A^t$, we already found:

$$\ker(A^t) = \ker(A) = \operatorname{span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}$$

Compute

$$\begin{pmatrix}1 & -1 & 1\end{pmatrix}\begin{pmatrix}2\\3\\1\end{pmatrix} = 1\cdot 2 + (-1)\cdot 3 + 1\cdot 1 = 2-3+1 = 0$$

Therefore $b \in \operatorname{im} A$.

(v) Let the columns of A be v_1, v_2, v_3. Vectors in $\operatorname{span} v_1 \cap \operatorname{span}\{v_2, v_3\}$ are of the form $x_1 v_1$ such that there are $x_2, x_3 \in F$ with $x_1 v_1 = x_2 v_2 + x_3 v_3$. That is $(x_1, -x_2, -x_3)^t$ is a solution to $Ax = 0$. We know that $\ker A = \{c(1,-1,1)^t : c \in F\}$ and hence $(x_1, -x_2, -x_3)^t = c(1,-1,1)^t$. So we have $\operatorname{span} v_1 \cap \operatorname{span}\{v_2, v_3\} = \{cv_1 : c \in F\}$.

Proof of Theorem 7.3. Let $A = (v_1 \ldots v_n)$ and R be an RREF obtained from A. Reordering the columns if necessary, we assume the pivot columns of R occur in positions 1 to r then $R = \begin{pmatrix} I_r & C \\ 0 & 0\end{pmatrix}$ where $C = (c_{ij}) \in M_{r\times(n-r)}(F)$.

(i) We are going to show that $v_1, \ldots, v_r$ is a basis of $\operatorname{im} A$. Let $x = (x_1, \ldots, x_r, 0, \ldots, 0)^t$. Suppose $x_1 v_1 + \cdots + x_r v_r = Ax = 0$. Then $0 = PAx = Rx = (x_1, \ldots, x_r, 0, \ldots, 0)^t$. Thus $x_1 = \cdots = x_r = 0$ and $v_1, \ldots, v_r$ are linearly independent. We write $R = (w_1 \ldots w_n)$. By the structure of RREF, for any $r+1 \le j \le n$ $w_j = \sum_{i=1}^r c_{ij} w_i$. Thus $x = (-c_{1j}, \ldots, -c_{rj}, 0, \ldots, 1, \ldots, 0)^t$ where 1 appears at the j-th component satisfies $Rx = 0$. Hence $Ax = 0$ and $v_j = \sum_{i=1}^r c_{ij} v_i$. This shows $\operatorname{span}(v_1, \ldots, v_r) = \operatorname{im} A$.

(ii) For a general solution of $Ax = 0$, solving $Rx = 0$ gives $x_i = -\sum_{j=r+1}^n c_{ij} x_j$ for $1 \le i \le r$. Hence

$$x = \begin{pmatrix} -\sum_{j=r+1}^n c_{1j} x_j \\ \vdots \\ -\sum_{j=r+1}^n c_{rj} x_j \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = x_{r+1}\begin{pmatrix} -c_{1(r+1)} \\ \vdots \\ -c_{r(r+1)} \\ 1 \\ \vdots \\ 0\end{pmatrix} + \cdots + x_n \begin{pmatrix} -c_{1n} \\ \vdots \\ -c_{rn} \\ 0 \\ \vdots \\ 1\end{pmatrix}.$$

For $r+1 \le j \le n$, let $u_j = (-c_{1j}, \ldots, -c_{rj}, 0, \ldots, 1, \ldots, 0)^t$. Then $u_j \in \ker A$ by construction. We know that $\ker A = \operatorname{span}(u_{j+1}, \ldots, u_n)$. Suppose $x_{r+1}u_{r+1} + \cdots + x_n u_n = 0$. By construction, last $n-r$ coordinates of x are $(x_{r+1}, \ldots, x_n)^t$. So $x_{r+1}u_{r+1} + \cdots + x_n u_n = 0$ implies $x_{r+1} = \cdots = x_n = 0$. Thus, the vectors are linearly independent. Therefore, $(u_{r+1}, \ldots, u_n)$ is a basis for $\ker A$.

(iii) Since $(PA)^t = A^t P^t$, we have $\operatorname{im}(R^t) = \operatorname{im}(A^t P^t) \subset \operatorname{im} A^t$. On the other hand $\operatorname{im}(A^t) = \operatorname{im}(R^t (P^{-1})^t) \subset \operatorname{im}(R^t)$. Hence $\operatorname{im} A^t = \operatorname{im} R^t$. For $1 \le i \le r$, we define $\varphi_i = (0, \ldots, 1, \ldots, 0, c_{i(r+1)}, \ldots, c_{in})^t$. Then $R^t = (\varphi_1 \ldots \varphi_r\, 0 \ldots 0)$. If $y_1, \ldots, y_r \in F$ are such that $y = y_1\varphi_1 + \cdots + y_r\varphi_r = 0$. Since the first r coordinates of y is $(y_1, \ldots, y_r)^t$, the equation implies $y_1 = \cdots = y_r = 0$. Since other rows of R are 0, the first r rows of R is a basis of $\operatorname{im} R^t = \operatorname{im} A^t$.

(iv) Suppose R and R' are two RREF of A. Then $R' = P'A$ and $R = PA$ where P and P' are products of elementary matrices. Then $R' = QR$ and where $Q = P'P^{-1}$

is invertible. As above, we assume $R = \begin{pmatrix} I_r & C \\ 0 & 0 \end{pmatrix}$ where $C = (c_{ij}) \in M_{r\times(n-r)}(F)$. We write $Q = \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix}$ where $Q_1 \in M_{r\times r}(F)$, $Q_2 \in M_{r\times(n-r)}(F)$, $Q_3 \in M_{(n-r)\times r}$ and $Q_4 \in M_{(n-r)\times(n-r)}(F)$. Then $R' = QR = \begin{pmatrix} Q_1 & Q_1C \\ Q_3 & Q_3C \end{pmatrix}$. By (i), the number of pivots is $\operatorname{rank} A = r$. Thus by definition of RREF, the last $m-r$ rows of R' are 0. Hence $Q_3 = 0$. We claim that Q_1 is invertible. Indeed, if this is not the case, then the first r columns of $Q = \begin{pmatrix} Q_1 & Q_2 \\ 0 & Q_4 \end{pmatrix}$ are linearly dependent. This contradicts the fact that Q is invertible. Thus Q_1 is invertible. Since $R' = \begin{pmatrix} Q_1 & Q_1C \\ 0 & 0 \end{pmatrix}$ is in RREF and the first r columns are linearly independent, they must all be the pivot columns since any non-pivot column is a linear combination of the pivot columns to the left of it. Thus using the definition of RREF, $Q_1 = I_r$. Hence $R' = R$. □

Corollary 7.9. *Let $A \in M_{n\times n}(F)$ be invertible. Then there exists elementary matrices $E_1, \dots, E_m$ such that $A = E_1 \cdots E_m$.*

Proof. Since A is invertible, $\operatorname{rank} A = n$. Then by Theorem 7.3, every column of A is a pivot column. Since pivot column in an RREF consist of 1 at the pivot and 0 elsewhere, the RREF for A is I. Since the Gauss Elmination is multiplying elementary matrices $M_1, \dots, M_m$ to the left of A, we have $M_m \cdots M_1 A = I$ and hence $A = E_1 \cdots E_m$ where $E_i = M_i^{-1}$ is also an elementary matrix. □

Corollary 7.10. *Compute A^{-1} using Gauss elimination:*

1. *Start with the augmented matrix $[A \mid I]$*
2. *Perform row operations to reduce the left side A to the identity matrix.*
3. *The result will be $[I \mid A^{-1}]$ where the right block becomes the inverse of A.*

Computationally, row operations correspond to multiplying by elementary matrices, so the whole process is: $E_k \cdots E_2 E_1 A = I \implies A^{-1} = E_k \cdots E_2 E_1$.

If at any step we cannot reduce A to the identity (e.g., a pivot is zero and cannot be fixed by swapping rows), then A is not invertible.

Proof of Theorem 7.6. If the last column of $[A \mid b]$ is a pivot, then the last nonzero rows of the RREF of $[A \mid b]$ contains a row of the form $(0 \quad 0 \quad \cdots \quad 0 \mid 1)$. This corresponds to the equation $0x_1 + 0x_2 + \cdots + 0x_n = 1$ which has no solution.

If the last column in the RREF of $[A \mid b]$ is not a pivot column, then the pivot appears in columns of A. Hence a basis of $\operatorname{im}[A \mid b]$ is a basis of $\operatorname{im} A$ i.e. $\operatorname{im}[A \mid b] = \operatorname{im} A$. Hence $b = Ax$ for some $x \in F^n$.

To find all solutions to $Ax = b$ we write the RREF of $[A \mid b]$ as $\left(\begin{array}{cc|c} I_r & (c_{ij}) & \tilde{b} \\ 0 & 0 & 0 \end{array}\right)$ after permuting columns. We solve the pivots as $x_i = \tilde{b}_i + \sum_{i=r+1}^{n} c_{ij} x_j$ for $1 \le i \le r$.

Then

$$x = \begin{pmatrix} \tilde{b}_1 - \sum_{j=r+1}^n c_{1j}x_j \\ \vdots \\ \tilde{b}_r - \sum_{j=r+1}^n c_{rj}x_j \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_{r+1} \begin{pmatrix} -c_{1(r+1)} \\ \vdots \\ -c_{r(r+1)} \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} -c_{1n} \\ \vdots \\ -c_{rn} \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

□

Remark 7.11. What if $b \notin \operatorname{im} A$ but we want to find the x such that Ax is closest to b? You would need the least square method which is covered in Linear Algebra II.

7.2 Determinant

Let $v_1, \ldots, v_3 \in \mathbb{R}^3$. We write $vol(v_1, v_2, v_3)$ to be the volume of the parallelepiped spanned by v_1, v_2, v_3. The *signed volume* of v_1, v_2, v_3, denoted by $\det(v_1, v_2, v_3)$ is a function that gives $vol(v_1, v_2, v_3)$ if v_1, v_2, v_3 are right handed and $-vol(v_1, v_2, v_3)$ if v_1, v_2, v_3 are left handed. The determinant is the n-dimensional generalization of this which satisfies the properties in the following definition.

Definition 7.12. The **determinant** is a function $\det : M_{n\times n}(F) \to F$ satisfying the following properties. We write $A = (v_1 \, \ldots \, v_n)$ where $v_i \in F^n$ are the columns of A.

(i) det is multilinear in columns: If for some $1 \leq j \leq n$, $v_j = ax + by$ where $a, b \in F$ $x, y \in F^n$, then

$$\det(v_1, \ldots, ax + by, \ldots, v_n) = a \det(v_1, \ldots, x, \ldots, v_n) + b \det(v_1, \ldots, y, \ldots, v_n).$$

(ii) det alternative in column: If $v_i = v_j$ for $i \neq j$, then $\det(v_1 \, \ldots \, v_n) = 0$.

(iii) $\det I = 1$.

Lemma 7.13. *Suppose* $\det : M_{n\times n}(F) \to F$ *is a function satisfying (i), (ii), (iii) of Definition 7.12. Then we have the following*

(i) $\det(v_1 \, \ldots \, 0 \, \ldots \, v_n) = 0$.

(ii) $\det(\ldots v_i \, \ldots \, v_j \ldots) = -\det(\ldots v_j \, \ldots \, v_i \ldots)$.

(iii)
$$\det(v_1 \, \ldots \, v_j + \sum_{i\neq j} \lambda_i v_i \, \ldots \, v_n) = \det(v_1 \, \ldots \, v_j \, \ldots \, v_n)$$

Proof. (i) By multilinearity, $\det(v_1 \, \ldots \, 0 \, \ldots \, v_n) = 0 \det(v_1 \, \ldots \, 0 \, \ldots \, v_n) = 0$.

(ii)

$$
\begin{aligned}
&\det(\ldots, v_i \ldots v_j, \ldots) + \det(\ldots, v_j \ldots v_i, \ldots)\\
=&\det(\ldots, v_i \ldots v_j, \ldots) + \det(\ldots, v_i \ldots v_i, \ldots)\\
&+\det(\ldots, v_j \ldots v_j, \ldots) + \det(\ldots, v_j \ldots v_i, \ldots)\\
=&\det(\ldots, v_i \ldots v_i + v_j, \ldots) + \det(\ldots, v_j \ldots v_i + v_j, \ldots)\\
=&\det(\ldots, v_i + v_j \ldots v_i + v_j, \ldots)\\
=&0.
\end{aligned}
$$

(iii) By multilinearity in the j-th column, we can expand:

$$\det(v_1 \ldots v_j + \sum_{i\neq j} \lambda_i v_i \ldots v_n) = \det(v_1 \ldots v_j \ldots v_n) + \sum_{i\neq j} \lambda_i \det(v_1 \ldots v_i \ldots v_n),$$

where in each term of the sum, the j-th column has been replaced by v_i. By the alternating property, each determinant in the sum vanishes, since two columns are equal in each: $\det(\ldots, v_i \ldots v_i, \ldots) = 0$. Therefore,

$$\det(v_1 \ldots v_j + \sum_{i\neq j} \lambda_i v_i \ldots v_n) = \det(v_1 \ldots v_j \ldots v_n)$$

$\square$